Novel Bioinformatics Tool: Interpretation of Glycan Mass Spectra with Metal Adducts and Multiple Adduct Combinations

Julian Saba[1], Ningombam Sanjib Meitei[2] and Arun Apte[2]

[1]Thermo Fisher Scientific, San Jose, CA, USA; [2]PREMIER Biosoft International, Palo Alto, CA, USA

Overview

Purpose: To demonstrate the use of SimGlycan software for interpretation of glycan mass spectra with metal adducts and multiple adduct combinations.

Methods: MS/MS and MS^n spectra of permethylated bovine fetuin and human IgG glycans were acquired on a Velos Pro ion trap mass spectrometer. Structural elucidation was performed using SimGlycan software.

Results: The combination of permethylation, MS^n and SimGlycan enabled structural Interpretation of glycan mass spectra with metal adducts and multiple adduct combinations.

Introduction

Mass spectrometry (MS) has emerged as a powerful tool for the structural elucidation of glycans. However, there are some drawbacks to using MS-based approaches. Mass spectrometers generate large volumes of data. Currently, processing of data from glycans is mostly done manually, which makes it tedious and time-consuming. Previously we had presented a bioinformatics tool for automated structural interpretation of glycan MS/MS and MS^n data.[1] This was limited to characterizing glycans with single adduct of H and Na ($[M+H]^{+1}$, $[M-H]^{-1}$, $[M+2H]^{+2}$, ($[M+2Na]^{+2}$ etc.). However, the ionization of glycans by MS results in the formation of several different adducts. These adducts maybe present as single adduct or combination of multiple adducts. Here we expand our bioinformatics tool, SimGlycan, to support lithium (Li), potassium (K) and magnesium (Mg) adducts as well as combination of multiple adducts such as Na + H, Li + H, Na+K etc. In order to demonstrate the utility of the software, a combination of permethylation and MS^n are used to characterize glycans derived from bovine fetuin and human IgG.

Methods

Sample Preparation

Bovine fetuin and human IgG (1 mg, Sigma) were reduced, alkylated and digested overnight with trypsin (Thermo Scientific, Rockford, IL, USA) in 25 mM ammonium bicarbonate buffer (pH~8) at 37 ºC. PNGase F solution (3 µL, Roche) was added to 200 µL of digested sample and the mixture was incubated for another 16 hours at 37 ºC. The released glycans were separated from the peptides using a Sep-Pak® C18 cartridge (Waters). The Sep-Pak C18 was conditioned by washing with acetonitrile, followed by water. PNGaseF digested sample was loaded onto the cartridge and the released glycans were eluted with 1% ethanol while the peptides remained bound to the Sep-Pak C18. The released glycans were first purified using a porous graphite carbon column (PhyNexus) and then permethylated as described previously.[2]

Mass Spectrometry

All MS experiments were performed using a Thermo Scientific Velos Pro dual pressure linear ion trap mass spectrometer via direct infusion into the nanoelectrospray source. The mass spectrometer settings and SimGlycan software version 4.0 (PREMIER Biosoft International) search parameters are listed in Tables 1 and 2.

Table 1. Mass Spectrometer Settings

Source	nano-ESI	Isolation Width	3
Capillary Temperature	200 ºC	Collision Energy	30
S-lens RF Level	50 %	Activation Time	10 ms
Source voltage [kV]	1.3	Predictive AGC Enabled	Yes
Full MS Mass Range	150-2000 (*m/z*)	No. Microscans for Full MS	5
Scan Rate	Enhanced	Target Value Full MS	3e4
Maximum Inj. Time Full MS	50 ms	Target Value MS^n	3e4
MS^n	50 ms		

Table 2. SimGlycan 4.0 Search Parameters

Ion Mode	Positive	Class	Glycoprotein
Adducts	Multiple	SubClass	N-Glycan
Precursor *m/z* Error Tolerance	0.8 Da	Biological Source	bov fetuin/h-IgG
Spectrum *m/z* Error Tolerance	0.8 Da	Pathway	Unknown
Chemical Derivatization	Permethylated	Search Structure	All
Reducing Terminal	Free	Glycan Type	All

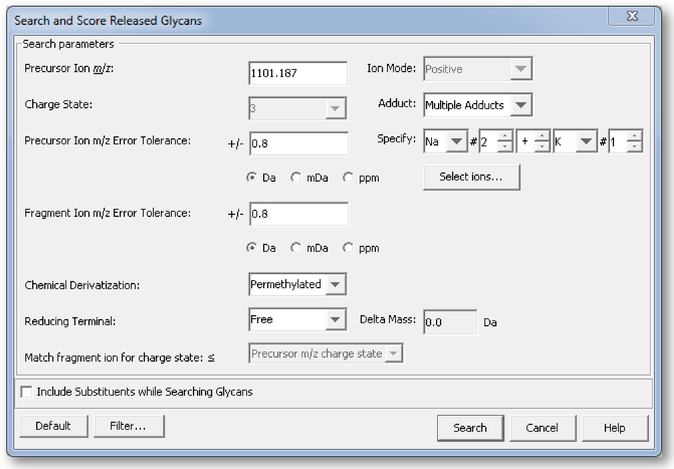
Results

Studies have shown that glycans are very susceptible to the effects of salts and other compounds. In most cases small amounts of sodium and alkali metals are added to improve ionization efficiency of glycans. However, the introduction of these adducts can result in spectra with precursors containing adducts other than H and in some cases multiple adducts in different combinations. These factors can complicate spectra interpretation as one must account for all these possible combinations at both the MS and MS/MS levels. Figure 1 shows a screen capture of the new SImGlycan search tab. Options for selecting adducts is shown on the top right hand side of the figure. The user can now select the types of adducts to search the spectra. Additionally, the user can specify either single adducts such as H, Na, Li, K or in combinations such as Na + H, Li + H, Na+K etc.

In order to test the performance of the software to handle multiple adducts as well as adducts other than H or Na, glycans released from bovine fetuin were chosen. This was an ideal system to test the capability of the software because the glycan content of bovine fetuin has been characterized in depth.

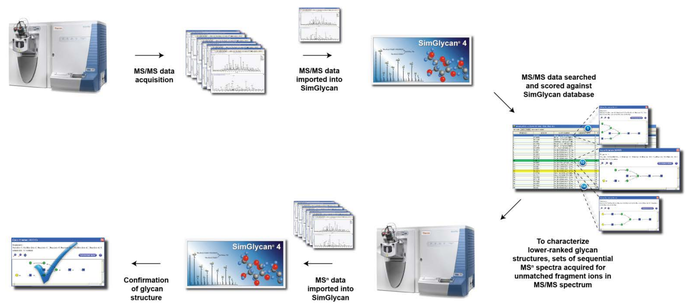
Figure 2 shows the overall workflow undertaken for these sets of experiments. Briefly, MS/MS spectral fragmentation data is acquired on Velos Pro mass spectrometer and then imported into SimGlycan software. SimGlycan automatically matches the experimentally acquired mass spectra against its comprehensive database of theoretical glycan fragments, and generates a list of candidate glycan structures. The SimGlycan database is a relational database containing 22,456 glycans, 22,814 glycoproteins, 11,438 glycans with known biological sources,

FIGURE 1. SimGlycan's search parameter window



11,918 glycans with known classes, 263 biochemical reactions, 194 biochemical pathways, 250 glycan related enzymes and 22,265 other database links. Each proposed glycan structure is assigned a rank and a score to reflect how closely it matches the experimental data. The rank is based on calculating the proximity score, which is a numerical representation of how closely the experimental properties of the glycan, such as composition and branching pattern, match with those of the glycans in the database. For these experiments the Velos Pro mass spectrometer was operated in "Enhanced Scan" profile mode allowing for charge state determination (up to 3+) of precursors and fragment ions. Manual examination of the MS profile showed that majority of the glycan precursors contained adducts other than H and in combination of multiple adducts. For example, a glycan observed at *m/z* 1101.36 contained both Na and K adducts.

FIGURE 2. Workflow for automated structural interpretation of MS^n glycan spectra.



For example, a glycan observed at *m/z* 1101.36 contained both Na and K adducts in the combination $[M+2Na+K]^{+3}$. In order to test the performance of the software, this particular precursor was targeted for MS/MS experiments. Data were imported into SimGlycan software for structural characterization. In cases where structural isomer differentiation was needed sequential MS^n spectra were acquired. SimGlycan characterized glycans were verified using manual assignment and previously published data.

Figure 3 shows an example of the results obtained at this stage of the workflow – at the MS/MS level. Prior to the recent implementation of multiple adduct support, data interpretation was done manually thus making it very tedious and time consuming. Additionally with the presence of multiple adducts the fragment ions can contain various combination of adducts as shown in Figure 4. Which further adds to the complexity of data interpretation.

Though structural interpretation can be performed using MS/MS-level spectral data, it's very difficult, if not impossible to determine structural isomers without MS^n-level spectral data. Examination of the glycan list (Figure 3) generated by SimGlycan software for the glycan at *m/z* 1101.36 revealed additional glycan compositions having identical mass which were scored much lower. Furthermore, we also observed for the top ranked glycan that when the experimental spectrum was mapped onto the theoretical spectrum, peaks showing high abundance remained unexplained (Figure 4). Though reported to have a much lower probability of matching the MS/MS spectrum, the lower ranked glycans could be additional isomers present in the sample.

To characterize a lower-ranked glycan structure, sets of sequential MS^n spectra are acquired for unmatched fragment ions. Each successive level of MS^n fragmentation spectra is then brought into SimGlycan software to compare the experimental and predicted MS^n fragmentation pathway. The software generates an annotated spectrum depicting the fragmentation matches and loss of consecutive monosaccharide units. As the level of MS^n increases, the fragments generated become increasingly structure specific and thus aid in characterizing specific structures, isomers and branching patterns.[2] Using sequential MS^n we were able to confirm additional structural isomers ranked 3 and 5 in Figure 3.

Figure 5 shows an additional example for a glycan observed at *m/z* 1221.58. This glycan similar to the one observed at m/z 1101.38 contains precursor with multiple adducts, $[M+2Na+K]^{+3}$ as well as the fragments that contains various combination of adducts.

We further extended this to glycans released from human IgG. This is an area of interest as IgGs are involved in human circulation as part of the humoral immune response and the changes in *N*-glycosylation of IgG associate with various diseases and affect the activity of therapeutic antibodies and intravenous immunoglobulins.

FIGURE 3. SimGlycan search results for ion trap MS/MS spectrum of precursor ion at *m/z* 1101.36 (+2). Symbolic representation of top-ranked glycans search results by SimGlycan software is shown.

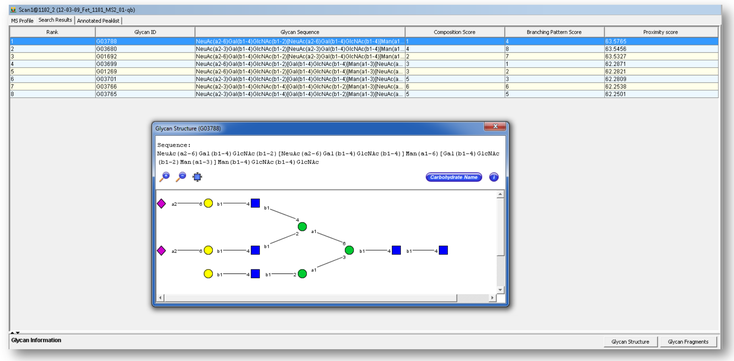


FIGURE 4. Ion trap MS/MS spectrum of permethylated bovine fetuin released glycan at *m/z* 1101.36. Spectrum annotated using symbolic representation.

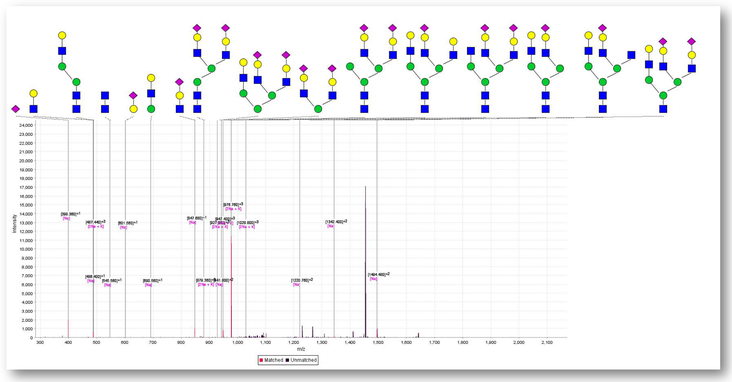
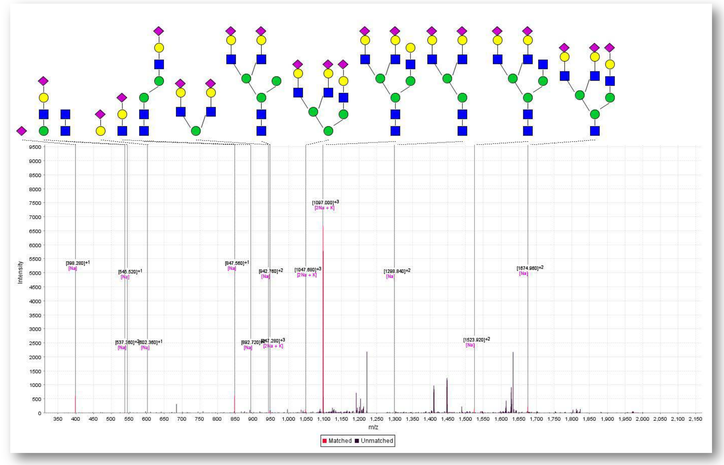


FIGURE 5. Ion trap MS/MS spectrum of permethylated bovine fetuin released glycan at *m/z* 1101.36. Spectrum annotated using symbolic representation.



Using a combination of infusion, permethylation and MS^n more than 20 structures of the complex type were identified for this human IgG glycoprotein.

Conclusion

- SimGlycan software simplifies data analysis by providing comprehensive support for MS experiments performed on Thermo Scientific ion trap and ion trap- Orbitrap hybrid mass spectrometers.
- SimGlycan has been expanded to provide support Li and K adducts as well as combination of multiple adducts such as Na + H, Li + H, Na+K etc.
- The overall analysis time was reduced to matter of minutes thus enabling truly automated, high-throughput data analysis.

References

1. Saba, J.; Apte, A.; Meitei, N.S.; Viner, R., Application Note 516: Automated Glycan Structural Isomer Differentiation Using SimGlycan Software .
2. Ciucanu, I.; Kerek, F., A Simple and Rapid Method for the Permethylation of Carbohydrates. *Carbohydrate Research* **1984,** 131, (2), 209-217.

Thermo SCIENTIFIC